ABELL EMEC VMAX ALL FLASH 950F OVERVIEW FOR MAINFRAME ENVIRONMENTS

ABSTRACT

This white paper describes the features that are available for DELL EMC® VMAX™ for IBM z systems. The latest Mainframe features are available with Mainframe Enabler V8.0 and higher. Throughout this document, VMAX refers to VMAX All Flash 950F.

July 2017
# TABLE OF CONTENTS

## EXECUTIVE SUMMARY

**AUDIENCE** ................................................................. 4

## INTRODUCTION

**VMAX ALL FLASH** .......................................................... 5

**HYPERMAX OS** .............................................................. 6

**Embedded Hypervisor** .................................................... 6

**Data at Rest Encryption (D@RE)** .......................................... 6

## MANAGEMENT SOFTWARE

**Mainframe Enabler** .......................................................... 7

**GDDR – Geographically Dispersed Disaster Restart** .................... 7

**Unisphere** .......................................................................... 7

## FACTORY-CONFIGURED ARRAYS

**Virtual Provisioning in VMAX 950, Pre-configured Arrays, Storage Resource Pools** ......................... 7

**Track Size** ........................................................................ 8

**TimeFinder SnapVX** .......................................................... 8

**zDP™ – Data Protector for z systems** ............................................ 9

**Enhanced SRDF** ................................................................. 10

**Support for FBA Data Transfer Over FICON** ........................... 11

## HARDWARE ENHANCEMENTS

**VMAX Engines** ............................................................... 11

**Multi-Core Emulation: Processing Power Where it’s Needed Most** ...................................................... 12

**Dynamic Virtual Matrix/Infiniband Fabric** ................................ 13

**Director Emulations** .......................................................... 13

**Vault to FLASH** ................................................................. 14

**6 Gb/s SAS back-end/drive infrastructure** .................................... 14

**16 Gb/s FICON and zHPF support** .......................................... 14

**IBM z Systems Compatibility Enhancements** .......................... 15

**Ultra dense Disk Array Enclosure (DAE) support** ....................... 15

**Local RAID: Performance and Physical Configuration Benefits** ....................................................... 16

**Dense Single Cabinet Configurations** ...................................... 16

**Bay (rack) Dispersion** ........................................................ 16
EXECUTIVE SUMMARY

Organizations around the globe need IT infrastructures that can deliver instant, continuous access to the massively increasing volumes of data associated with traditional online transaction, batch processing, and big data use cases such as data warehousing and data analytics. This must be accomplished with a continuous reduction in TCO, improvement in Storage Service Level Agreements (SLAs) and mitigation of risk associated with storing the data. Many are contractually bound to SLAs that describe required levels of service, often with penalties associated with non-compliance. Organizations are trying to understand how the new generation of 'systems of engagement' applications, built around the world of social, mobile, cloud, and big data (collectively named the “3rd Platform” by IDC) can be leveraged on the mainframe (known as the 1st platform) which serves as the “system of record” for most large organizations. New threats to data availability and integrity are surfacing almost weekly and IT organizations must respond with state-of-the-art techniques to protect their data. DELL EMC has been helping enterprise solve mainframe storage problems for decades and is now redefining the traditional storage array, morphing it into a “Data Services platform” that will become the bridge between the 1st and 3rd platforms to modernize and deliver the next generation of Hybrid Cloud computing and storage with the ultimate in availability, data integrity, and TCO management.

Modern mainframe storage architectures require:

- Massive capacity scaling
- Massive performance scaling
- Flexibility to handle highly fluctuating workloads yet maintain consistent service levels, all the time
- Both physical and logical protection from threats to data integrity and availability
- A data protection infrastructure that can scale across an unlimited number of arrays, sysplexes, and physical locations
- Reduced costs through infrastructure hyper-convergence
- A usage model that is automated and almost totally hands off, unifying management with other platforms and further reducing TCO

The mainframe-enabled VMAX 950F array is designed to meet and exceed these requirements through:

- Scaling to 480 drives in a single two engine cabinet
- Leveraging a scale out architecture up to 8 engines, 576 cores, 16TB cache and 256 physical host connections
- Leveraging the powerful Dynamic Virtual Matrix architecture in a single storage tier.
- The most advanced local and remote replication solutions in the industry
- A converged storage platform capable of running powerful storage and application workloads on VMAX including Mainframe, Open, IBM i
- An ease of use model that is unique in the high end storage arena

This paper explains how the VMAX delivers these capabilities and more for mainframe environments.

AUDIENCE

This white paper is intended for DELL EMC customers and those evaluating DELL EMC storage for purchase.

INTRODUCTION

DELL EMC’s VMAX is incredibly well positioned to solve the CIO’s challenges of embracing a modernized, flash-centric data center while simultaneously trying to simplify, automate, and consolidate IT operations. VMAX isn’t just bigger, better, and faster (although it is!), VMAX is a flexible Data Services platform that specifically addresses the new requirements of the modern mainframe data center while continuing to deliver the reliability and availability our customers have relied on for years.
With VMAX, the industry’s leading tier 1 array has evolved into a thin-provisioned hardware platform with a complete set of rich software data services including Data Protector for z systems (zDP), a revolutionary new data protection solution that enables rapid recovery from logical data corruption, be it from simple processing errors or malicious intent. VMAX data services are delivered by a highly resilient, agile hardware platform that offers global cache, CPU (processing) flexibility, performance and high availability at scale to meet the most demanding storage requirements.

VMAX All Flash also radically simplifies management at scale by eliminating the need for storage administrators to answer the questions “how many disks of which type to allocate” and “where does my data need to be placed?”. Tier 1 storage management can now be done in a matter of minutes, and doesn’t require extensively trained IT storage administrators. By delivering these capabilities, VMAX improves overall staff productivity, giving them time to focus on the needs of the business, rather than management of the technology.

DELL EMC VMAX arrays continue the mainframe legacy of all Symmetrix, DMX, and VMAX arrays that have come before it. They enable IT departments to consolidate, economize and reduce risk within their data center infrastructure while delivering mission-critical storage with scale, performance, availability, security and agility from DELL EMC that companies have relied on for years.

**VMAX ALL FLASH 950F**

The world’s most powerful All Flash system is the new VMAX 950F array, shown in Figure 1 below. The 950F is based on the newer 18 core Intel® Broadwell processor and can simultaneously support both mainframe and Open Systems workloads to achieve even greater TCO benefits.

VMAX 905F for mainframe:

- Supports mainframe, open systems, IBM i, block and file storage with six 9’s availability
- Implements simplified ‘zBrick’ packaging that combines a VMAX engine, 2 x DAEs, and 13TB of storage capacity. Additional capacity is implemented by adding 13TB Flash Packs called “zCapacity packs”
- Offer industry leading data services including advanced replication, storage management, data protection, access to hybrid cloud, and data encryption.
- Leverage the powerful Dynamic Virtual Matrix Architecture and HYPERMAX OS
- Leverage the latest 3.8TB, 1.9TB, and 960GB flash drives in a single configurable zCapacity Pack of 13TB usable.

**FIGURE 1: VMAX 950F ALL FLASH ARRAY**
HYPERMAX OS

The previous versions of the VMAX Family operating system were called Enginuity. Starting with VMAX, the internal operating system is called HYPERMAX OS, the industry's first open storage and hypervisor converged operating system. HYPERMAX OS combines industry-leading high availability, I/O management, Quality of Service (QoS), data integrity validation, storage tiering, and data security with an open application platform. It features the first real-time, non-disruptive storage hypervisor that manages and protects embedded services by extending VMAX high availability to services that traditionally would have run external to the array. It also provides direct access to hardware resources to maximize performance. The hypervisor can be non-disruptively upgraded.

The VMAX hypervisor reduces external hardware and networking requirements, delivers higher levels of availability, and dramatically lowers latency. HYPERMAX OS runs the Dynamic Virtual Matrix, leveraging its scale-out flexibility of cores, cache, and host interfaces.

HYPERMAX OS provides:

- The management of system resources to intelligently optimize performance across a wide range of I/O requirements and ensure system availability through advanced fault monitoring, detection, and correction capabilities
- Concurrent maintenance and serviceability features
- The foundation for specific software features available through DELL EMC’s disaster recovery, business continuity, and storage management software
- Functional services for VMAX arrays and for a large suite of DELL EMC storage application software
- Automated task prioritization, including basic system maintenance, I/O processing, and application processing

Embedded Hypervisor

HYPERMAX OS derives its name from the inclusion of a hypervisor which enables embedded data services to execute directly on the storage array delivering new levels of efficiency to enterprise workloads. This guest operating system environment is currently used to provide these services:

1. Monitoring and control of a single VMAX via a ‘tools’ guest hosting Solutions Enabler, SMI-S, and Unisphere for array management and performance monitoring
2. The analytics components of Fully Automated Storage Tiering pattern recognition as well as host-provided hint translation services. This component is used by a number of internal VMAX technologies to determine the priority of a transaction.
3. Embedded NAS (eNAS) which provides flexible and secure multi-protocol file sharing (NFS, CIFS/SMB 3.0) as well as multiple file server identities (CIFS and NFS servers)

Data at Rest Encryption (D@RE)

Data in enterprise storage must be secure, both inside and outside of the VMAX. D@RE (Data at Rest Encryption) ensures that the potential exposure of sensitive data on discarded, misplaced, or stolen media is reduced or eliminated. D@RE provides hardware-based, on-array, back-end encryption for VMAX models running HYPERMAX OS. Encryption within each individual disk drive (“back-end” encryption) protects your information from unauthorized access even when drives are removed from the system. D@RE provides encryption on the back end using SAS I/O modules that incorporate XTS-AES 256-bit data-at-rest encryption. These modules encrypt and decrypt data as it is being written to or read from a drive. All configured drives are encrypted, including spares. In addition, all array data is encrypted, including Symmetrix File System and Vault contents.

D@RE incorporates RSA® Embedded Key Manager for key management which provides a separate, unique DEK (Device Encryption Key) for all drives in the array including spare drives. D@RE keys are self-managed, so there is no need to replicate keys across volume snapshots or remote sites.
As long as the key used to encrypt the data is secured, encrypted data cannot be read. In addition to protecting against threats related to physical removal of media, this also means that media can readily be repurposed by destroying the encryption key used for securing the data previously stored on that media. Customers can now also leverage secure external key managers for data at rest encryption via the KMIP standard.

D@RE is compatible with all VMAX system features, allows for encryption of any supported logical drive types or volume emulations and delivers powerful encryption without performance degradation or disruption to existing applications or infrastructure.

**MANAGEMENT SOFTWARE**

**Mainframe Enabler**

Mainframe Enabler (MFE) is a suite of z/OS based products for managing your VMAX in a z/OS environment. MFE commands can be used to monitor device configuration and status and perform control operations on devices and data objects within your DELL EMC VMAX storage environment.

Mainframe Enabler 8.0 or above is required for VMAX arrays running the HYPERMAX OS 5977 releases. MFE 8.0 is also downward compatible with older VMAX and Symmetrix arrays.

**GDDR – Geographically Dispersed Disaster Restart**

Geographically Dispersed Disaster Restart (GDDR) is Dell EMC’s automated continuity product for both primary (disk) and backup (tape) storage. It provides automated recovery from all manner of disaster and also provides unplanned swap capability. GDDR monitors the environment and provides alerts to potential problems, and automates their resolution. GDDR was the first in the industry to provide 3-site and 4-site solutions. GDDR can automate both mainframe and open systems environments together, and synchronize recovery of data across platforms. GDDR has been enhanced to support VMAX All Flash 950. Version 5.0 includes support for the new TimeFinder SnapVX local replication function, described below.

**Unisphere**

Unisphere enables customers to easily provision, manage, and monitor VMAX environments. Unisphere 8.1 has been enhanced to support the new capabilities of the VMAX All Flash 950.

With HYPERMAX OS, it is possible to run Unisphere for VMAX as a Guest Operating system directly within the VMAX native hypervisor, eliminating the need for an external management host, and associated fibre channel adapters to control and manage the VMAX array in a FICON attached environment. The Embedded Management option is defaulted when ordering the VMAX 950 system as CPU and memory requirements must be sized appropriately. Please see the Unisphere for VMAX Documentation available at [https://support.emc.com](https://support.emc.com) for more information.

Unisphere offers simple “big-button” navigation and streamlines operations to simplify and reduce the time required to manage VMAX; it also simplifies storage management under a common framework.

Unisphere for VMAX contains a number of task-oriented dashboards to make monitoring and configuring VMAX systems intuitive and easy. As an example, the Storage Group Dashboard displays information about application storage groups and whether or not they are meeting their SLO requirements. Administrators can quickly navigate from this dashboard to gather more in-depth performance statistics.

**FACTORY-CONFIGURED ARRAYS**

**Virtual Provisioning in VMAX 950, Pre-configured Arrays, Storage Resource Pools**

All VMAX 950 arrays arrive pre-configured from the factory with Virtual Provisioning Pools ready for use. VMAX pools all its drives into a Storage Resource Pool (SRP) which provides physical storage for thin devices that are presented to hosts. VMAX All Flash arrays require no initial setup by the storage administrator, reducing the time to I/O and radically simplifying the management of VMAX storage. With the SRP, capacity is monitored at the
SRP level and disk pools, RAID levels, and thin device (TDEV) binding are no longer constructs the storage administrator needs to manage. All thin devices are ready for use upon creation and RAID is implemented under the covers in the SRP as part of the pre-configuration. Figure 2 shows the SRP components and the relationship to the storage group (SG) used for grouping thin devices to the host applications. Note there is a 1:1 relationship between disk groups and data pools. Each disk groups specifies a RAID protection and disk size forming the basis for each of the preconfigured thin pools. Every VMAX array comes from the factory with the bin file (configuration) already created. This means best practices for deployment - TDAT sizes, RAID protection, and data pools - will already be in place and no longer have to be created or managed by the storage administrator.

![Figure 2: Storage Resource Pool Components](image)

With the new preconfigured SRP model, VMAX provides all the benefits of Thin Provisioning without the complexity.


**Track Size**

VMAX arrays support 3380 and 3390 CKD volumes using a single 56KB track size as the allocation unit for storage from SRPs for thin provisioned devices. When thin devices are created in a VMAX array they consume no space from the SRP until they are first written.

**TimeFinder SnapVX**

EMC TimeFinder® software delivers point-in-time copies of volumes that can be used for backups, testing, data recovery, database system cloning, data warehouse refreshes, or any other process that requires parallel access to production data.

HYPERMAX OS 5977 for VMAX introduced TimeFinder SnapVX, which combines the best parts of the previous TimeFinder offerings with new ease-of-use features, increased scalability and significantly improved space efficiency.

In arrays running HYPERMAX OS, TimeFinder SnapVX lets you non-disruptively create point-in-time copies (snapshots) of critical data at the volume level. SnapVX creates snapshots by storing pre-update images of tracks (snapshot deltas) directly in the SRP of the source device. These “point in time” snapshots only consume
space when source tracks are updated. Tracks that are not updated share allocations across many snapshots, enabling the creation of many point in time copies of a volume without consuming additional space. SnapVX is also a “targetless snapshot” design, meaning a target volume is not required to obtain a point in time copy of a volume. In other words, the capture of a “point in time” has been separated from its use. Therefore, with SnapVX, you do not need to specify a target device and source/target pairs when you create a snapshot. If there is ever a need for the application to use the point-in-time data, you create links from the snapshot to one or more target devices. If there are multiple snapshots and the application needs to find a particular point-in-time copy for host access, you can link and re-link until the correct snapshot is located.

SnapVX for CKD volumes supports TimeFinder/Clone, TimeFinder/Snap (virtual devices), and TimeFinder/Mirror via emulations that transparently convert these legacy TimeFinder commands to SnapVX commands. You can still run jobs that use TimeFinder/Clone, TimeFinder/Snap, and TimeFinder/Mirror commands, but the underlying mechanism within HYPERMAX OS is SnapVX.

In HYPERMAX OS arrays, SnapVX supports up to 256 snapshots per source device (including any emulation mode snapshots). Legacy session limits still apply to the emulations of prior TimeFinder offerings. SnapVX and legacy TimeFinder operations, as well as Flashcopy emulation, can only coexist on source volumes. Intermixing these technologies across source and target volumes is not supported at this time.

You can set snapshots to automatically terminate after a specified number of days or at a specified date and time. Figure 3 shows multiple snapshots of a production volume with a Time to Live (TTL) of one day. HYPERMAX OS will only terminate the snapshot if it does not have any links to target volumes. If it does have links, HYPERMAX OS will terminate the snapshot when the last link has been unlinked. Writes to a linked target device will only be applied to the linked target and will not change the point in time of the snapshot itself. Snaps can be deleted in any order without affecting their sequencing.

For more information refer to Tech Note VMAX Local Replication Suite TimeFinder SnapVX and TimeFinder Emulation

zDP™ – Data Protector for z Systems

Much of the focus on data protection in the last twenty years has been on recovery from loss of a data center due to unplanned outages or disasters. The emphasis has been on providing copies of data at alternate sites and on ensuring that data integrity of the copies is preserved. Availability with data integrity has been the goal.

In recent years there has been an alarming number of examples of data corruption due to processing errors or malicious actors that result not in a loss of data availability, but a loss of data integrity in the production environment. All the storage-based replication technology deployed to protect against loss of data since the invention of data replication, provides no protection at all against data corruption, and in fact dutifully replicates corrupted data to all recovery sites with impressive speed and accuracy!
With data corruption risk taking on new and more dangerous forms beyond processing errors that, at best, introduce errant data to the more serious willful hacking and destruction of data, the responsibility of CIOs has expanded beyond rapid recovery from data center loss to rapid recovery from loss of data integrity.

Data Protector for z Systems (zDP) is designed to address the problem of large scale recovery from logical corruption. zDP is an DELL EMC industry exclusive z/OS-based application that utilizes SnapVX snapshots to enable rapid recovery from logical data corruption. zDP achieves this by providing multiple, frequent, and consistent point-in-time copies of data in an automated fashion across multiple volumes from which an application level recovery can be conducted. By providing easy access to multiple different point-in-time copies of data (with a granularity of minutes), precise remediation of logical data corruption can be performed using storage or application-based recovery procedures. zDP provides the following benefits:

- Faster recovery times as less data must be processed due to the granularity of the available point in time data copies
- Cross application data consistency for recovery data
- Minimal data loss compared to the previous method of restoring data from daily or weekly backups.

This is especially important for non-DBMS data, which does not have the granular recovery options provided by log files and image copies associated with database management systems.

Prior to zDP, the only way to recover from logical data corruption was an offline copy, either a BCV (Business Continuance Volume), sometimes known as a “Gold Copy” or a backup made to offline physical or virtual tape. Even in the best datacenters practicing the latest data protection procedures, often only one offline copy of the “state of the business” was being made per day. Considering that 144 Snapshots can be taken in a 24 hour period (at 10 minute intervals) with zDP as compared to a single BCV or offline tape backup, zDP gives you 144x the granularity to recover from a situation that could have otherwise been detrimental or fatal to your business.

zDP Enhancements
Mainframe Enablers and Hypermax code further enhanced zDP’s usability and monitoring capabilities. New features as of May 2017 include:

1) The QUERY SNAPSET command with the SNAPSET option provides a new entry in the output called ‘RDP Cache Utilization’ percentage. This information allows customers to monitor their VMAX Replication Data Pointers (RDP) usage.

2) Three new parameters on the VDG which are available in zDP ISPF Create VDG panels or batch JCL statements.
   a. RDP Cache Util(ww,cc) – ‘ww’ is a percentage for when warning messages should be sent to z/OS. ‘cc’ is a percentage when the RDP threshold reaches a critical level and messages will be sent to z/OS.
   b. SMF record number – This permits the customer to set a SMF record number for zDP. SMF records are cut for that VDG at the beginning of each cycle.
   c. TRACKS – Provides additional output for your VDG for all total changed and unique tracks within the VDG.

3) TERMINATE RANGE command which provides the ability for customers to terminate multiple zDP snapsets with one statement. This command is available via zDP ISPF panels and batch JCL statements.

4) If a zDP snapshot fails during the ACTIVATE command, zDP will automatically clean up and terminate the snapsets.

Enhanced SRDF
The Symmetrix Remote Data Facility (SRDF) family of software is the gold standard for remote replication in mission critical environments. Built for the industry-leading high-end VMAX hardware architecture, the SRDF
family of solutions has been trusted for disaster recovery and business continuity for more than two decades. Asynchronous SRDF (SRDF/A) enables remote data services to provide 6 9s of data availability (31.5 seconds of system downtime a year) and 24x7xForever operation. Synchronous SRDF (SRDF/S) can achieve 7 9s of availability (3.2 seconds of downtime a year). The SRDF family offers unmatched deployment flexibility and massive scalability to deliver a wide range of distance replication capabilities.

Another key change in VMAX is an enhancement to the DSE (Delta Set Extension) feature which is designed to increase availability in SRDF/A environments. There is no longer a need to configure a separate pool in the array and there is no need for a DSE pool to exist in the remote (R2) array. Instead, the SRP will have a maximum DSE capacity associated with it (specified in GBs). DSE capacity is specified when the VMAX array is configured, resulting in a less complex configuration for the storage administrator to manage.

SRDF/A has also been improved to provide better resiliency and shorter and more predictable Recovery Point Objectives (RPOs) when operating under stress. This is done through an enhancement called Multi-cycle Mode (MCM) which allows more than two delta sets to exist on the source array. MCM has the benefit of always cycle switching based on the user set cycle time, so the cycles are now predictable and much smaller when applied to the secondary side. This eliminates the need for DSE on the secondary side.

SRDF/S has also been enhanced to provide increased performance of devices through reductions in replication processing overhead.

In addition, since all VMAX directors are capable of supporting a variable number of ports (up to 16 ports for every director configured with emulation for SRDF) the number of SRDF groups supported on an individual VMAX SRDF director (RA) has increased from 64 to 250, which is also the total number of SRDF groups allowed per VMAX array.

All VMAX array models are also capable of supporting enhanced hardware compression for bandwidth optimization on both IP and fibre links.

**Support for FBA Data Transfer Over FICON**

VMAX All Flash 950 also supports the storage of non-mainframe data (fixed block) being transferred over FICON. This was added after HYPERMAX OS 5977 in support of applications by vendors that enable open systems data storage on the array in a z/OS mainframe environment.

**HARDWARE ENHANCEMENTS**

**VMAX Engines**

The VMAX hardware infrastructure for all models is designed around the concept of “engines”. A VMAX engine is a pair of physically separate director boards housed in the same enclosure. In order to enhance availability, each director board is physically independent, with its own Intel® processors, power feed and redundant hardware. The VMAX 950F director board supports Intel Broadwell processors with up to 18 physical cores each (50% more than previous models). An 8 engine VMAX 950F can employ up to 576 Intel processor cores. Since the HyperMAX OS exploits the SMT2 feature of the Intel processors, in reality there are up to 1152 logical processor cores deployed in a VMAX 950F!

Each VMAX engine was designed to be modular and redundant for ease of service and elimination of any offline outage. Directors can be removed from the front of the rack for service upgrades without the need to disconnect any cabling from front-end or back-end I/O modules. The only physical differences between the engines across the various VMAX models is the configuration of the dual inline memory modules (DIMMs) and the number and operating frequency of the CPU cores.

All VMAX arrays contain two Management Module Control Systems (MMCS) in system bay 1. This helps to increase system availability as there are multiple access points to the system for remote access. If there is a failure in either MMCS, the system is able to dial home from the remaining MMCS for remote recovery or diagnose if hardware replacement is required. The MMCS replaces the Service Processor that was present in earlier VMAX models. Figure 4 below shows the common hardware components of the VMAX engines as well as
the Management Module Control Station (MMCS), which is only required in the first system bay of each VMAX array.

**Multi-Core Emulation: Processing Power Where it’s Needed Most**

VMAX arrays can be configured to allocate more CPU cores to handle host I/O, allocate more CPU cores to handle disk I/O, or allocate CPU cores evenly between front and back end operations. Pre-defined CPU core mappings allow specification of performance characteristics based on expected I/O profiles and usage of the system. Most, but not all, mainframe workloads require front-end centric configurations and are so configured in the factory.

This flexibility is made possible by Multi-Core emulation which improves the CPU and physical port utilization of HYPERMAX OS, extending the proven VMAX code architecture while improving overall performance.

Figure 5 shows the default Multi-Core emulation in VMAX arrays. Cores are pooled for front end, back end, and for HYPERMAX OS functions. Multiple CPU cores on the director will work on I/O from all of the ports. This helps ensure VMAX directors’ ports are always balanced.
Dynamic Virtual Matrix/Infiniband Fabric

The Dynamic Virtual Matrix provides the Global Memory interface between directors with more than one engine. The Dynamic Virtual Matrix is composed of multiple elements, including Infiniband Host Channel Adapter (HCA) endpoints, Infiniband Interconnects (switches), and high-speed passive, active copper, and optical serial cables to provide a Virtual Matrix interconnect.

A fabric Application Specific Integrated Circuit (ASIC) switch resides within a special Management Interface Board Enclosure (MIBE), which is responsible for Virtual Matrix initialization and management. Each fabric port connects back to an Infiniband switch housed in the first system bay cabinet. The Infiniband switches are only present in multi-engine systems and are added with the addition of a second engine. Infiniband switches are installed in pairs and each director has a path to Fabric switch A and B. Fabric switches are supported by standby power supplies for vault activities to ensure all cache data gets vaulted.

The VMAX 950 array supports 16 Interconnect ports. This is shown below in Figure 6.

![FIGURE 6: 18 PORT INFINIBAND FABRIC SWITCH FOR THE VMAX 950](image)

The cabling of the Infiniband switches is simpler than previous VMAX models, enabling faster setup time by DELL EMC field personnel.

Director Emulations

As previously discussed, the VMAX director consists of sets of hardware modules or boards which vary by model – the VMAX 950 contains Intel Broadwell CPUs. Director boards are configured into engines. Within a VMAX system, the directors work together to apply their resources to the work of many low-level functions called “emulations,” 10 of which are named in the table below.

Two new director emulations have been introduced with HYPERMAX OS 5977 on VMAX: Infrastructure Manager (IM) and HYPERMAX OS Data Services emulation. The IM emulation is an aggregation of common infrastructure tasks previously distributed across all director types. This consolidation is intended to allow other directors to devote their CPU resources to I/O specific work only, without interference from the demands of the infrastructure tasks. The HYPERMAX OS Data Services emulation also provides a consolidation of various functionalities, with the main goals being to both reduce I/O path latency and introduce better scalability for various HYPERMAX OS applications.

<table>
<thead>
<tr>
<th>Emulation</th>
<th>Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>Back</td>
<td>Disk Services (DA for SAS Disks)</td>
</tr>
<tr>
<td></td>
<td>End</td>
<td></td>
</tr>
<tr>
<td>DX</td>
<td>Back</td>
<td>Director External (Used for FAST.X and ProtectPoint)</td>
</tr>
<tr>
<td></td>
<td>End</td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>Middle</td>
<td>Infrastructure Management</td>
</tr>
<tr>
<td>EDS</td>
<td>Middle</td>
<td>Enginuity Data Services</td>
</tr>
<tr>
<td>EF</td>
<td>Front</td>
<td>Front End (FICON)</td>
</tr>
<tr>
<td></td>
<td>End</td>
<td></td>
</tr>
<tr>
<td>FA</td>
<td>Front</td>
<td>Front End (Fibre Channel)</td>
</tr>
<tr>
<td></td>
<td>End</td>
<td></td>
</tr>
</tbody>
</table>

1 May also be referred to as EDS or ED.
Vault to FLASH

Vaulting is the process of saving Global Memory data to a reserved space within the VMAX during an offline event. Vault to FLASH provides vaulting of Global Memory data to internal flash I/O module(s). This feature provides the following advantages:

- Improved array performance due to larger Global Memory per director, capable of being saved within 5 minutes
- Physically, the VMAX weighs less (fewer batteries are required to save data when a power interruption is detected)
- VMAX is easier to configure as there is no longer a requirement to reserve capacity on back-end drives for vault space
- A minimum drive count per engine is no longer required

6 Gb/s SAS back-end/drive infrastructure

All VMAX models utilize 6 Gb/s SAS (Serial Attached SCSI) drives with a “back-end” configuration that provides improved performance over legacy architectures. SAS is a high-speed, extremely reliable protocol that uses the same low-level technology as Fibre Channel encoding. SAS topology is different from Fibre Channel as SAS uses a connectionless tree structure with unique paths to individual devices. Routing tables store these paths and help to route I/O to the required locations.

16 Gb/s FICON and zHPF support

The VMAX 950 supports 16 Gb/s FICON. Each FICON channel adapter card (SLIC) within the VMAX consists of a 4 port 16 Gb/s I/O module based on the industry-standard QLogic chip set that auto-negotiates with the host to support 4, 8, and 16 Gb/s link speeds. It is possible to configure up to 32 16 Gb/s FICON ports for each VMAX engine.

VMAX is 100% zHPF compatible. This includes:

- List Prefetch and Bi-Directional support. These features enable a single I/O to efficiently access contiguous extents on a volume. This results in improved performance; for example, in DB2 when accessing indexes with poor cluster ratios (disorganized index scans).
- Format Write commands. This capability improves performance of utilities, such as DB2 load, reorg, index rebuilds and restores by enabling channel programs employing format writes to deliver large amounts of data in a single I/O.
- Exploitation of zHPF by the BSAM, QSAM, and BPAM access methods. In addition VMAX offers support for the following FICON enhancements announced with the IBM z13:
- Forward Error Correction (FEC) support for 16Gb/s FICON. This feature improves control over transmission errors on noisy fibre links and allows FICON to operate at higher speed over longer distances
- FICON Dynamic Routing (FIDR) which allows FICON to use dynamic routing policies for Inter-Switch Links (ISL) in the SAN
- Read Diagnostic Parameters to enable SAN management products to display diagnostic data for 16Gb/s links.
- zHPF Extended Distance II FICON which allows large write operations (> 64 KB) at distances up to 100 km to be executed in a single round trip, providing up to 50% I/O service time improvement and benefitting GDPS HyperSwap configurations

**IBM z Systems Compatibility Enhancements**

VMAX has added support for the following IBM 2107, copy services and other features:

- Query Host Access (QHA) CCW support. QHA is a used to determine if a device has active FICON path groups and is exploited by several host applications including the ICKDSF utility parameter VERIFY OFFLINE to check for LPAR access before initializing a volume.
- PPRC Soft Fence which prevents users from accidentally accessing the original PPRC primary volumes after a Hyperswap or PPRC Primary failure occurs.
- SPID Fencing which prevents any hosts from bringing a device or group of devices online
- Non-Disruptive State Save (NDSS). NDSS is intended for capturing diagnostic information on demand within the VMAX when certain problems occur in GDPS/PPRC environments
- IBM zHyperWrite. This feature is implemented within the PPRC (Metro Mirror) support mode of VMAX. It is exploited by DB2 in a Hyperswap enabled environment to write directly to the PPRC primary and secondary volumes that contain the active log datasets, bypassing PPRC replication for write I/Os and improving response time for this performance critical component of DB2.
- Dynamic Volume Expansion (DVE) which enables increasing the size of a logical volume while it is online to a host system (available for Open System and IBM System z™ hosts)
- Extended Address Volumes (EAVs) up to 1 TB are supported vs. the previous 223GB EAV device definition. EAV volumes are of interest to customers facing UCB constraints, consolidations, and growth (particularly in conjunction with DVE capability to grow volumes dynamically).

**Ultra dense Disk Array Enclosure (DAE) support**

VMAX 950 systems provide an ultra-high density DAE supporting up to 120 2.5” drives each. Figure 7 shows the DAE.

DAEs can be added to systems in single increments if using RAID 1, RAID 5 (3+1), and RAID 6 (6+2). However, if your system contains RAID 5 (7+1) or RAID 6 (14+2), adding DAEs may only be possible in pairs. A VMAX engine is able to support up to 6 DAEs (720 x 2.5” drives).

When the system is configured at the factory, drives are distributed across engines in balanced configurations to provide the optimal performance in the array. When drives are added it is expected that they will also be added in a balanced manner.

Every DAE has 4 power zones and can thus continue to operate despite the loss of power to any one zone (which would require loss of two separate power supplies, also known as a double fault). If required, it is possible to configure RAID 6 (14+2) and RAID 5 (7+1) across 4 DAE so that only one member resides in any power zone.
Local RAID: Performance and Physical Configuration Benefits

VMAX arrays implement local RAID which requires all members of a RAID group to be associated with the same engine. This ensures local access and control over I/O for all RAID members and reduces the number of messages and Global Memory operations that need to be carried out for RAID operations, lowering I/O overhead and improving RAID performance.

Local RAID also eliminates the need for cross-bay (cross-frame) cabling in direct/daisy chain DAEs. This allows for the physical separation of a multiple frame VMAX system at the engine/bay level (in order to position the frames around any obstacles or across an aisle in the datacenter), making the VMAX systems the most flexible storage system in the industry.

Dense Single Cabinet Configurations

All VMAX arrays can be configured with a single engine per cabinet and up to 6 DAEs. Alternatively a system can be configured to have 2 engines per system bay with 4 DAEs (up to 480 2.5” drives) to provide a much denser storage configuration, up to 4.42 PBe. A single bay can support 64 host ports and up to 4TB of cache in a single standard floor tile. Figure 8 shows the layout configuration.

Bay (rack) Dispersion

VMAX 950 system racks can be physically separated by up to 25 meters to avoid columns and other obstacles in the data center without a need to reserve empty floor tiles for future array growth. Any VMAX system bay can be placed anywhere in your data center as long as it is within 82 feet (25 meters) of the first system bay which
houses the Infiniband Dynamic Virtual Matrix switches. Figure 9 shows a possible dispersion (separation) option for an 8-engine VMAX 950 with 2 adjacent system bays and 6 system bays dispersed at a distance of 25M each from system bay 1.

![Figure 9: BAY DISPERSION WITH VMAX](image)

**Third-Party Racking**

All VMAX arrays support industry-standard 19-inch racks and optional third-party racking for ease in conforming to your existing data center infrastructure. Third Party racks must meet the dimensions set out in the DELL EMC VMAX All Flash Site Planning Guide available on [https://support.emc.com](https://support.emc.com)

**CONCLUSION**

The industry's leading, field-proven, tier-1 array, VMAX, offers a complete set of rich software data services including a revolutionary new continuous data protection solution for mainframe users: zDP. VMAX data services are delivered by a highly resilient, agile hardware platform that offers global cache, CPU flexibility, performance and the most FICON ports of any array in the industry to satisfy the most demanding storage infrastructure needs, whether converged (both mainframe and non-mainframe workloads running within the same VMAX 950F) or not.

VMAX arrays are designed and built for management simplicity, extreme performance, hyper-consolidation and massive scalability in a dense footprint. DELL EMC introduced the storage industry’s first Dynamic Virtual Matrix and brought data services closer to the storage they access, eliminating the need to proliferate functionally-limited “data appliances” in the data center. The VMAX Data Services Platform enables flexible storage infrastructure decisions to be made that are not bound by what is capable within an appliance’s “frame.” This approach provides hyper consolidation, excellent Total Cost of Ownership (TCO), simple and agile management, while exceeding customers’ current and future needs for mainframe storage.

To learn more about how DELL EMC products, services, and solutions can help solve your business and IT challenges, contact your local representative or authorized reseller, visit [www.emc.com](http://www.emc.com), or explore and compare products in the [DELL EMC Store](http://www.emc.com).

updated h13805
GDDR for VMAX_wp
The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any DELL EMC software described in this publication requires an applicable software license.

EMC², EMC, the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners. © Copyright 2017 EMC Corporation. All rights reserved. Published in the USA. July 2017 White Paper Part Number H14909.7

EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

EMC is now part of the Dell group of companies